

chapter one
Overview and future of molecular
biomarkers of exposure and early disease
in environmental health

William A. Suk and Samuel H. Wilson

Contents

I. Introduction	4
II. History and definitions	5
III. Defining criteria for useful biomarkers	7
A. Accuracy and reproducibility	7
B. Specificity /sensitivity	8
C. Database retrieval and retrospective analysis	8
D. Experimental models and biological plausibility	8
E. Sampling requirements	9
F. High-throughput/feasibility on population-scale	9
C. Implementing biomarker criteria	9
IV. Assessment of current status of biomarker research	9
V. Fulfilling the promise of biomarker research: instruments for future progress	12
A. The biomarker validation process	12
B. Biomarker database development	13
C. Improved peer review of biomarker-related research proposals	13
D. Sample banks and repositories to support biomarker development	14
VI. Closing remarks	14
References	14

Abstract This chapter reviews the status of molecular biomarker research at the beginning of the 21st century. The characteristics of useful biomarkers are discussed and mechanisms to facilitate progress in biomarker research are proposed. These mechanisms are concerned with four important goals:

1. Developing and promoting a consistent, efficient and effective biomarker validation process
2. Establishing and maintaining large robust relational databases to support biomarker development and use
3. Promoting expert, informed peer review of research proposals related to biomarker development
4. Improving access to biological and/or environmental samples to support biomarker research and development

These mechanisms will help bring molecular biomarkers from the laboratory environment to the clinical/population-based setting, where they can have their intended impact on reducing the burden of human disease and protecting susceptible individuals from adverse and unnecessary risk.

1. Introduction

Validated molecular biomarkers have long been recognized as invaluable tools for identifying and preventing human disease. The potential of molecular biomarkers is especially high in relation to preventing environmentally-induced disease. This is an important focus of health research at present because of significant concern over the risk of human exposure to persistent organic pollutants, heavy metals, airborne pollutants, environmental estrogens, and other environmental agents.¹ Molecular biomarkers could play a key role in facilitating advances in disease detection and prevention; however, as discussed in this chapter, much work is needed before the great potential for biomarkers in environmental health is to be realized.

This chapter is an overview of the current status of molecular biomarker research at the beginning of the 21st century. We discuss the characteristics of useful biomarkers and define criteria that can be used during biomarker development. We also focus on mechanisms that could promote progress in biomarker research. These mechanisms are concerned with four important goals:

1. Developing and promoting a consistent, efficient and effective biomarker validation process
2. Establishing and maintaining large robust relational databases to support biomarker development and use
3. Promoting expert, informed peer review of research proposals related to biomarker development
4. Improving access to biological and/or environmental samples to support biomarker research and development

Together, efforts in these areas will facilitate more rapid development of biomarkers that can be effectively applied at the population scale. As biomarkers begin to be applied more widely, it is also important to assure that they are implemented ethically, with attention to the social and legal issues associated with their use. Future success in biomarker development and validation will undoubtedly enhance our ability to understand, treat, and prevent environmentally-induced disease. Widespread ethical application of validated biomarkers will bring great benefit to public health.

II. History and definitions

Biomarkers have been used for some time in laboratories and clinics to monitor metabolic status, disease and/or toxic responses in tissues and cells from animals and humans. They have a long history of success in the fields of analytical chemistry and have been implemented and promoted by the National Institute of Standards and Technology (NIST) and the National Center of Environmental Health at the Centers for Disease Control and Prevention (CDC). In the 1980s, rapid advances were made in molecular approaches to biology, genetics, biochemistry, and medicine, and many scientists recognized that these methods could facilitate development of molecular biomarkers. Molecular biomarkers appeared to hold the promise of transforming toxicology, epidemiology, environmental health, and clinical medicine leading to increased molecular understanding of disease and valuable applications in molecular epidemiology. With the emergence of new technologies in the last decade, including genomics-based approaches, the field of biomarker development and application has assumed even greater importance. Ultimately, biomarkers offer the promise to improve disease prevention and reduce burden of exposure and disease in the human population worldwide.

To develop molecular biomarkers in environmental health, scientists distinct and distant fields must come together and learn to work in an interdisciplinary manner. The expertise of epidemiologists, toxicologists, statisticians, chemists, biologists, pathologists, and other scientists is needed. Because these scientists work in different ways utilizing different tools, methods, and languages, it can take considerable time for productive collaborations to develop. To address this issue, conferences have been held to promote interactions between diverse groups of scientists interested in biomarkers and to foster interdisciplinary research related to biomarker development. The National Institute of Environmental Health Sciences (NIEHS) has played a role in this process, sponsoring the landmark 1990 conference "Application of Molecular Biomarkers in Epidemiology."² This conference was designed to increase the visibility of research opportunities in environmental epidemiology. More recently, NIEHS sponsored the "International Conference on Arctic Development, Pollution and Biomarkers of Human Health" in Anchorage, Alaska, during which a broad assessment was made of the state of biomarker research and its impact on research on the health of the Arctic environment.³ The present volume is also an assess-

ment of progress in biomarker research and is in part an extension of the efforts initiated at the NIEHS-sponsored conference in Anchorage, Alaska.

A molecular paradigm or framework for biomarker development was presented in the 1983 report of the National Research Council on Risk Assessment in the Federal Government.' This paradigm describes progression along a continuum from exposure to disease that is influenced by the following variables: external dose, internal dose, target tissue dose, metabolic activation, detoxification and early and late biological effects (see Chapter 21 by J. Groopman, Figure 1 or Chapter 25 by Greenlee, Figure 1). Recent iterations of this paradigm⁴ recognize that progression from exposure to disease can also be influenced by preventive interventions and genetic and/or environmental susceptibility of exposed individuals.

The National Academy of Sciences defines "biomarker" as an "indicator signaling events in biological systems or samples." In the context of this chapter, a biomarker is a measurement of a molecular or chemical substance or event in a biological system. Nonmolecular biomarkers have also been developed and are used extensively to measure behavioral and/or cognitive effects in humans and animals. These biomarkers are useful, but they are outside the scope of this discussion, which concerns itself with molecular biomarkers in environmental health, research and medicine.

There are three broad types of molecular biomarkers in the field of environmental health: biomarkers of exposure, biomarkers of effect, and biomarkers of susceptibility.-' Biomarkers of exposure are methods that quantify body burden of chemicals or metabolites and they are usually applied early in the exposure-disease pathway. These markers are powerful tools for epidemiologists, allowing relatively accurate measurement of external and/or internal dose of an environmental agent; in contrast, when biomarkers of exposure are not available, epidemiologists often rely on much less quantitative approaches or mathematical models to extrapolate from external to internal dose. The kinetics of exposure complicate use of exposure biomarkers, so they should be used with attention to and awareness of this issue.

Biomarkers of effect detect functional change in the biological system under study and allow investigators to predict the outcome of exposure. However, an exposure to an environmental agent often does not have a lasting impact on a biological system. Early markers of effect can be used to identify populations at risk of disease or toxicity; late markers of effect are more tightly linked to disease progression. Ultimately, biomarkers of exposure and effect are used in a coordinated fashion to understand and prevent disease progression.

Individual susceptibility to exposure or disease is influenced by multiple genetic and environmental factors, including genetic polymorphism, age, pre-existing disease, diet, occupation, behavior and lifestyle.' Differential susceptibility to some diseases is also associated with ethnicity and gender. Recent advances in molecular biology, genomics, toxicology, and in understanding disease mechanisms can be used to identify susceptible subpopulations. The Human Genome Project and the Environmental Genome Project

have already had a huge impact on understanding genetic susceptibility to exposure and disease, and much more information on this topic will become available in the next few years. Polymorphisms are rapidly being identified including those in human genes involved in activation and metabolism of xenobiotic agents (e.g., glutathione-S-transferase, N-acetyl transferase). These polymorphisms can be used as biomarkers of susceptibility. Gene and protein expression technology (DNA microarray and proteomics) and highthroughput genotyping methods will facilitate development of biomarkers of susceptibility. Understanding the complex interplay between genes and environment is a tremendous challenge to scientists. Nevertheless, identification of human disease susceptibility genes is a key step towards progress in reducing the burden of human disease.

Rapid progress in genomics and genomics-associated technology has brought increased awareness of the ethical, legal, and social issues (ELSI) related to scientific research and the use or misuse of biomedical information. These issues are highly relevant to molecular biomarkers and their application, and they have received significant attention in this field. Potential misuse of biomarker data can lead to denial of employment, denial of insurance, or social marginalization, and without appropriate patient counseling, biomarker studies can adversely affect psychological welfare, long-term life goals, and family planning. Many of these have been discussed elsewhere extensively⁶ (see Chapter 2 by Sharp and Zigas). Managing ELSI is an integral part of developing biomarkers for environmental health research.

III. Defining criteria for useful biomarkers

The following discussion cites many of the important characteristics of a useful molecular biomarker, many of which have been discussed previously^{5,7} and are discussed elsewhere in this volume. It may be unrealistic to expect that every biomarker will fulfill all the criteria; in that case, judicious use of several biomarkers that provide complementing and confirming data is a suitable alternative. This concept of using a suite of biomarkers was pioneered by Perera and colleagues,⁸ and it is often a beneficial approach. The strengths and/or deficiencies of each molecular biomarker should be carefully considered in experimental design.

A. Accuracy and reproducibility

A useful biomarker is highly accurate and reproducible, characteristics that may be thoroughly established during the development phase of the marker using statistical methods and laboratory studies. Internal controls are extremely valuable for ascertaining accuracy and reproducibility. Controls should also be used to characterize and quantify the dynamic range of the assay. Controls enhance data comparability between experiments in a single study, between distinct studies in a single lab, and between studies in different labs or with different models. Use of such controls can be built into the

data management structure as well. Attention to accuracy and reproducibility is important as assays move from laboratory studies with model systems to clinical or population-based studies with large numbers of samples.

B. Specificity/sensitivity

Biomarkers are powerful tools because they can identify and quantify exposure, effect, or susceptibility in individual members of a population. Thus, a biomarker must be sufficiently sensitive to provide an accurate measurement in a sample of limited quantity from a single individual. Sensitivity and specificity are often inversely related to one another. Specificity must be sufficient to avoid a high false positive rate, and sensitivity must be sufficient to avoid a high false negative rate. Standards and controls can be used to analyze and optimize assay sensitivity and specificity.

C. Database retrieval and retrospective analysis

Data management and data comparability are extremely important in development and validation of biomarkers and for their use in population-based studies. Proper data management allows researchers to use data efficiently and to increase the data's value. Ideally, the database structure should allow the data to be reevaluated by third parties long after they are added to the database. For this reason, accessions to the database should include complete sets of raw data, and the database structure should be relational. Thus, researchers will be able to reexamine data when new insights are gained into disease processes (e.g., disease-related covariables are identified). Importantly, the statistical power of any study may increase significantly, when data from different studies are combined in a valid manner. With appropriate foresight, relational database structures can be optimized with these goals in mind.

D. Experimental models and biological plausibility

Biomarkers of exposure and biomarkers of effect distribute along the continuum from exposure to disease and in some cases, a biomarker can indicate both types of events (e.g., site-specific DNA adduct in the p53 gene and/or site-specific p53 mutation as marker of exposure to cigarette smoke and as early marker of lung cancer¹²). In most cases, the link between exposure and effect is ambiguous unless it is rigorously confirmed using animal and cellular models. Animal models allow researchers to understand the mode of action of environmental chemicals and the biological mechanism of their action. Such information leads to biologically plausible biomarkers of the disease process. In addition, cell and animal models are essential for developing and refining knowledge of dose-response and exposure kinetics. Thus, the biological basis and usefulness of a molecular biomarker can be enhanced by thorough use of cell and animal models.

E. Sampling requirements

Useful molecular biomarkers for population-based studies should rely on noninvasive sampling (i.e., buccal swab, urine, blood). This requirement reflects the need for high-throughput biomarkers and for large sample size in epidemiological studies. Sample banking is also important because it enhances access to experimental materials and facilitates high-quality research studies and retrospective analyses. Retrospective analyses can take advantage of newly developed technology or allow a new hypothesis to be tested without additional field work.

F. High-throughput/feasibility on population scale

Biomarker assays need high throughput capability for application in population-based studies. Current technological advances facilitate this requirement. DNA microarray, proteomics, immunodetection, and high-throughput sequencing or genotyping all have the potential for rapid analysis of many samples and endpoints using automated processing and data analysis.

G. Implementing biomarker criteria

It will be important to promote development of peer-reviewed guidelines that emphasize and/or require that a biomarker meets an accepted set of criteria and will be implemented in an ethical manner. These guidelines need to be implemented during the development phase of a biomarker as well as after the biomarker is implemented in a clinical or population-based setting. Mechanisms to fulfill this need have been presented previously^{7,13} and are discussed below. A recent report discussing guidelines for development of biomarkers of cancer¹⁴ recommended a phased program for biomarker development, akin to phases of drug development.

IV. Assessment of current status of biomarker research

It has been almost two decades since a molecular biomarker paradigm for study of environmental disease was advanced and discussed by the National Academy of Sciences.^{1,14,15} Much research has gone into developing biomarkers of exposure and effect that help monitor, treat, and/or prevent disease. It is reasonable at this time to begin to assess progress in molecular biomarker research and to evaluate how well this field has fulfilled its promise to enhance environmental health and to help identify, treat and/or prevent environmentally induced disease.

Only a few molecular biomarkers are currently in use in a clinical/population setting. Among the most notable used in clinical medicine are biomarkers for exposure to aflatoxin and lead. However, in a broad sense, it is fair to say that the promise of biomarkers has not yet moved from the lab to the clinic or from development to application. For example, the recent

conference "Arctic Development, Pollution, and Biomarkers of Human Health: May 1-3, 2000" produced a set of recommendations and a summary document that indicate that "assessment of specific exposure using biomarker technologies holds great promise... [but] these techniques are not immediately available for widespread use." A similar conclusion was expressed in 2000 by Bennett and Waters." Some researchers also express concern about the limitations of biomarkers in epidemiological studies.¹⁷

Nevertheless, there are exemplary results with biomarkers that can be mentioned. A recent summary highlighted the success achieved with biomarkers of exposure to aflatoxin.⁷ A biologically plausible model based on adverse effects of aflatoxin bioadducts supports the strong epidemiological link between adduct level and risk of liver cancer. In both animal models and human studies, DNA adduct level correlates with known or estimated external dose of aflatoxin. A urinary metabolite of the aflatoxin DNA adduct has also been used effectively as a quantitative biomarker of exposure. In addition, large-scale epidemiological studies (cross-sectional, longitudinal, and case-control) have been carried out using these biomarkers. Prospective studies are also ongoing and biomarkers of aflatoxin exposure have been used to monitor efficacy of intervention during clinical trials of the drug oltipraz. The aflatoxin biomarker is a unique example of a well-developed and useful molecular biomarker of exposure; unfortunately, few other biomarkers have been implemented in a similar manner.

There are other promising exposure biomarkers that are based on detection of DNA and/or protein adducts (i.e., PAH-adducts including adducts from exposure to benzo-[a]pyrene, chromium-DNA adducts). However, several complications exist when using these and other DNA adducts as biomarkers. These complications include variable rates of carcinogen activation and adduct repair and the validity of using surrogate tissue in place of the target tissue in which the biological effects of a DNA lesion occur. Sufficient numbers of studies with animal models are needed to understand the dose-response curve for each environmental agent and to attempt to better define the dynamic range of the exposure-response relationship. This is especially important for evaluating chronic low-level exposure in humans.

Many studies have focused on assays for genetic instability based on the premise that genetic instability is associated with risk of cancer and certain other chronic disease and conditions. These studies propose to develop biomarkers using assays that detect DNA strand breaks, chromosomal aberrations, micronuclei, or the endogenous rate of DNA repair or mutation. In general, these assays are suboptimal for biomarker development, largely because of unresolved issues in their reproducibility, accuracy, validation and interpretation. Poor reproducibility is commonly reported in interindividual and intra-individual measurements using genomic instability assays. These assays are often labor- and cost-intensive, and are poorly suited to high-throughput applications in population-based studies. Lastly, these assays are most commonly carried out in lymphocytes as a surrogate for the

target tissue at risk for disease^{18,19}; the relevance of genetic stability measurements in this and other surrogate tissues has not been clearly established.

Some of the most promising areas of biomarker research involve recent advances in genomics and genomics-related fields. Toxicogenomics is an important emerging field that exploits the genomics-based technologies of DNA microarray and proteomics to solve classical problems in toxicology. One of the leading research programs in toxicogenomics is the NIEHS National Center for Toxicogenomics (NCT). Through the effort of the NCT and other groups, DNA microarray and proteomics are expected to provide useful molecular biomarkers that can be characterized and implemented rapidly. These efforts will undoubtedly also facilitate discovery of many new disease susceptibility alleles in the human genome.

Advances in human genomics have already led to identification of numerous disease-related DNA polymorphisms. Many of these polymorphisms have been characterized in clinical studies, and in some cases, human genetic polymorphism is strongly associated with disease risk and/or susceptibility to environmental agents (e.g., alleles of glutathione-S-transferase and N-acetyl transferase, mutation of p53 and other tumor suppressor genes). As new high-throughput technologies are developed for simultaneous analysis of multiple genes, many additional disease-related polymorphisms will be discovered. The usefulness and cost effectiveness of research on human genetic polymorphism will be enhanced if human haplotypes are also identified and studied. Increased attention and effort in this area is warranted.

Other new technologies are bringing powerful capabilities to the researcher. One of these emerging technologies is nanotechnology, which is the creation of functional materials, devices, and systems at the nanometer scale (i.e., 1 to 100 nanometers) and the exploitation of novel nanoscale materials and phenomena. In the same way that miniaturization has changed the world of electronics, nanotechnology has the potential to revolutionize the fields of environmental health, biomedicine, pharmaceuticals, and biotechnology. Nanoscale analytical tools could make it possible to characterize chemical and mechanical properties of cells and to discover novel processes and a wide range of tools, materials, devices, and systems with unique characteristics. Eventually, by coupling advances in the knowledge of living systems with the unique capabilities of nanostructures and materials, it may be possible to detect and intervene in disease states using biologically inspired solutions.

Nanotechnology has the potential to influence future developments in the field of molecular biomarkers. For example, nanotechnology could be important in developing biosensor technologies for detecting and analyzing molecular targets in blood, saliva, clinical specimens, and chemical or biological materials in the living body. Nanotechnology could also play a role in validation of biomarkers in population-based studies. Other important emerging capabilities include automated workstations using DNA microarrays that extract, amplify, hybridize, and detect DNA sequences, protein

chips for high-throughput proteomic analyses, nano-opticochemical sensors for real-time imaging of physiological processes, and nanoarrays. It is anticipated that successful development of these new approaches could facilitate rapid advances in developing sensitive, accurate, and useful molecular biomarkers. Nevertheless, these techniques and their applications are currently being developed and refined, and their anticipated impact on clinical medicine lies several years in the future.

V. Fulfilling the promise of biomarker research: instruments for future progress

Many approaches are available to enhance future progress in the process of molecular biomarker development. We have identified the following four goals as critical steps in this process:

1. Development of a consistent, efficient and effective biomarker validation process
2. Establishment of large robust relational databases to support biomarker development
3. Development of a mechanism for expert peer review of research proposals related to biomarker development
4. Establishment of large repositories of relevant biological and environmental samples

A. The biomarker validation process

Many scientists, clinicians, and health and environmental policy makers have discussed the need for validated biomarkers and for mechanisms to promote the process of biomarker validation. Systematic efforts toward improving the biomarker validation process are critical. Biomarker validation is essential because it can assure development of biomarkers that meet specified criteria, namely, biomarkers should be accurate, reproducible, specific, sensitive, biologically plausible, high-throughput, and appropriate for use in population studies and in databases.

Because concerted effort is needed to improve biomarker validation, it is appropriate to establish validation conferences and one or more expert validation committees to work towards this goal. These groups could codify the criteria for biomarker validation and recommend approaches to standardize biomarker development. Targeted consensus development conferences for specific biomarkers would be appropriate and useful.

Existing validation committees provide a model for the process of biomarker validation. One such validation committee and process is the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) established in 1997 by the NIEHS. ICCVAM is composed of representatives from 15 federal regulatory and research agencies that

generate, use, or provide information relating to toxicity test methods for risk assessment. ICCVAM establishes criteria and processes that facilitate development, validation, and regulatory acceptance of new test methods. Through the National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM), ICC VAM also carries out peer reviews and organizes workshops on toxicological test methods. Clearly, biomarker development would be facilitated by formation of a committee whose role is similar to ICCVAM but whose focus is on biomarker assays instead of toxicology assays.

B. Biomarker database development

As discussed above, the future of biomarker research and development will depend on the size and quality of the databases that are used to manage biomarker data. Assay standardization and comparability are essential. This can be promoted by building relational database structures that are accessible to and utilized by many research groups. A large relational database that includes the results of many different studies is important because it increases the value of existing data and saves research dollars and time. Importantly, quality control measures can also be integrated into the database structure, ensuring that data accepted into the database meet essential criteria. In a similar manner, databases can be used to establish research design standards and "best practices" and promote them throughout the research community.

Awareness of these issues is generally high in the emerging field of toxicogenomics, and the efforts underway to develop a microarray database are instructive. For example, the Microarray Gene Expression Database (MGED) group was initially established at the Microarray Gene Expression Database meeting in November 1999. As indicated on the MGED Web site (www.mged.org), "The goal of MGED is to facilitate the adoption* of standards for DNA-array experiment annotation and data representation, as well as the introduction of standard experimental controls and data normalization methods. The underlying goal is to facilitate the establishment of gene expression data repositories, comparability of gene expression data from different sources and interoperability of different gene expression databases and data analysis software." Similar approaches are highly appropriate in the field of molecular biomarkers. The challenges are perhaps greater in this field because biomarkers use many different molecular techniques, each of which presents distinct issues with regard to standardization and comparability.

C. Improved peer review of biomarker-related research proposals

Biomarker development is clearly an interdisciplinary endeavor that requires many types of scientific expertise. In addition, biomarker development and

validation are closely linked to population-based molecular epidemiology studies. It is well recognized that this type of research does not fit well into any one of the study sections used by the National Institutes of Health (NIH) for peer-review of research grants. Therefore, these proposals have an unusually low success rate. This is a major impediment to progress in biomarker research, and it is strongly recommended that a new peer-reviewed advisory group or study section be considered for proposals in the area of molecular epidemiology and biomarker development.

D. Sample banks and repositories to support biomarker development

Restricted sample size often limits experimental design in biomarker development and validation. If samples are archived and stored in a sample bank or repository, then sample size can potentially be increased by use of archival material. This opportunity should be exploited as much as possible in biomarker studies. Researchers should therefore be encouraged to establish and maintain sample repositories and to deposit appropriate materials in repositories whenever possible.

VI. Closing remarks

The end of the 20th century brought with it a revolution in molecular biology that culminated in advances such as completion of a preliminary draft of the human genome²⁰ and whole genome expression analyses using DNA microarray technology. These advances brought an optimism to the fields of toxicology and environmental health and the anticipation that molecular biomarkers might soon come of age and have a major impact in environmental health. Optimism is justified for the future of molecular biomarkers, but current progress in biomarker application has not fulfilled the expectations of many scientists and policy makers.

In this discussion, we present a challenge to researchers in all scientific fields relating to environmental health. This challenge is to develop mechanisms that allow the potential of molecular biomarkers to be fulfilled as rapidly as possible. These mechanisms are needed so that biomarkers are brought from the laboratory environment to the clinical- and populationbased arena, where they can have an impact on health issues that affect the human population. Ultimately, ethical application of validated biomarkers has the potential to significantly reduce burden of exposure and disease and to protect susceptible individuals from adverse and unnecessary risk. Concerted and focused efforts should allow that potential to be fulfilled, with great benefit to human health.

References

1. National Research Council, *Risk Assessment in the Federal Government: Managing the Process*. National Academy Press, Washington, D.C., 1983.

2. Smith, M. and Suk, W., Application of molecular biomarkers in epidemiology, *Environ. Health Perspect.*, 102 (Suppl. 1), 229, 1994.
3. Summary of the International Conference on Arctic Development, Pollution and Biomarkers of Human Health, National Institute of Environmental Health Sciences, 2001.
4. Olden, K. and Wilson, S., Environmental health and genomics: visions and implications, *Nat. Rev. Genet.*, 1, 149, 2000.
5. Committee on Biological Markers of National Research Council, Biological markers in environmental health research, *Environ. Health Perspect.*, 74,3, 1987.
6. Christiani, D.C. et al., Applying genomic technologies in environmental health research: challenges and opportunities, **1**. *Occup. Environ. Med.*, 43, 526, 2001.
7. Groopman, J.D. and Kensler, T.W., The light at the end of the tunnel for chemical-specific biomarkers: daylight or headlight?, *Carcinogenesis*, 20, 1, 1999.
8. Perera, F. and Weinstein, I.B., Molecular epidemiology: recent advances and future directions, *Carcinogenesis*, 21, 517, 2000.
9. Perera, F., The potential usefulness of biological markers in risk assessment, *Environ. Health Perspect.*, 76, 141, 1987.
10. Perera, F. and Weinstein, I.B., Molecular epidemiology and carcinogen-DNA adduct detection: new approaches to studies of human cancer causation, *J. Chronic Dis.*, 35, 581, 1982.
11. Hainaut, P. and Pfeifer, G.P., Patterns of p.53 C to T transversions in lung cancers reflect the primary mutagenic signature of DNA-damage by tobacco smoke, *Carcinogenesis*, 22, 367, 2001.
12. Greenblat, M.S. et al., Mutations in the p53 tumor suppressor gene: clues to cancer etiology and molecular pathogenesis, *Cancer Res.*, 55, 4855, 1994.
13. Sullivan, Pepe M. et al., Phases of biomarker development for early detection of cancer, *J. Natl. Cancer Inst.*, 93, 1054, 2001.
14. National Research Council, *Biologic Markers in Pulmonary Toxicology*, National Academy Press, Washington, D.C., 1989.
15. National Research Council, *Biologic Markers in Reproductive Toxicology*. National Academy Press, Washington, D.C., 1989.
16. Bennett, D.A. and Waters, M.D., Applying biomarker research, *Environ. Health Perspect.*, 108(9), 907-910, 2000.
17. Wild, C.P. et al., A critical evaluation of the application of biomarkers in epidemiological studies on diet and health, *Br. J. Nutr.*, 6, 37, 2001.
18. Bonassi, S. et al., Are chromosome aberrations in circulating lymphocytes predictive of a future cancer onset in humans? Preliminary results of an Italian cohort study, *Cancer Genet. Cytogenet.*, 79, 133, 1995.
19. Vogel, U. et al., DNA repair capacity: inconsistency between effect of overexpression of five NER genes and the correlation to mRNA levels in primary lymphocytes, *Mutat. Res.*, 461, 197, 2000.
20. Venter, J.C. et al., The sequence of the human genome, *Science*, 291, 1304, 2001.
21. McPherson, J.D. et al., A physical map of the human genome, *Nature*, 409, 934, 2001.
22. Roberts, C.J. et al., Signaling and circuitry of multiple MAPK pathways revealed by a matrix of global gene expression profiles, *Science*, 287, 873, 2000.